

Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma

Year: 2023 | Citations: 491 | Authors: Y. Yeo, Jamil S. Samaan, Wee Han Ng, Peng-Sheng Ting, H. Trivedi

Abstract

Background: Patients with cirrhosis and hepatocellular carcinoma (HCC) require extensive care. Personalized education can improve their outcomes. ChatGPT (Generative Pre-trained Transformer), a natural language processing model, has shown potential to provide professional yet patient-friendly responses. **Aim:** To examine the accuracy and reproducibility of ChatGPT in responding to questions regarding knowledge, management, and emotional support for cirrhosis and HCC. **Method:** ChatGPT's responses to 164 frequently asked questions were independently graded by two transplant hepatologists, with a third reviewer resolving any discrepancies. We also compared the performance of ChatGPT on two previously validated and published questionnaires to the physicians or trainees who were tested in the included publications. Furthermore, we formulated the 26 quality measures of cirrhosis management into questions and tested ChatGPT's knowledge in cirrhosis care. Finally, the capacity to provide emotional support to patients or caregivers was tested. **Results:** ChatGPT regurgitated extensive knowledge about both cirrhosis and HCC, but for questions with correct responses, only a small proportion was labelled as comprehensive. The performance was better in basic knowledge, lifestyle, and treatment than in the domains of diagnosis and preventive medicine. For the quality measures, the model answered 76.9% of questions correctly but failed to specify the cut-off values for making medical decisions and treatment durations. When compared to physicians/trainees, ChatGPT fell short in knowledge of guidelines varying across geographic regions, such as HCC screening criteria. The model also provided practical and multifaceted advice to patients and caregivers regarding the next steps and adjusting to a new diagnosis. **Conclusion:** In summary, we analyzed the areas of robustness and limitations of ChatGPT's responses on the management of cirrhosis and HCC and relevant emotional support. ChatGPT may have a role as an adjunct informational tool for patients and physicians to improve outcomes.